An outline of pulse code modulation

by C. BUDD, A7884*

IF it is required to send a signal over any transmission path there will be advantages, from the viewpoint of discriminating against random noise, in coding the signal into a series of binary numbers and transmitting those numbers as a pulse train (Fig 1) and then decoding the pulse train at the receiver to recover the original signal.

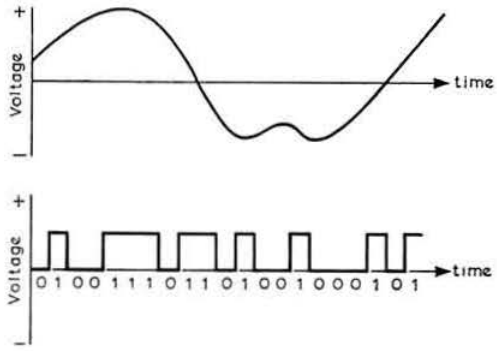

Fig 1. A possible signal waveform and its binary-coded equivalent

Suppose that the pulse train in Fig 1 was sent along the cable and that the waveform received at the other end was as shown in Fig 2, complete with noise and distortion. Now, in spite of the noise, it is quite possible to perceive from the waveform of Fig 2 the approximate periods during which pulses were and were not present. This information is all that is needed at the receiving end, provided that all the pulse-periods (ie periods during which pulses may or may not be present) are of equal length, to reconstruct the original pulse-train perfectly and hence to decode a noise-free and distortion-free copy of the original signal. If, on the other hand, the original signal had been sent along the cable just as it was, no amount of processing could have effectively removed the noise and distortion so introduced. Thus, unless the amplitude of the noise is at least comparable to that of the pulses, no significant impairment will result to the signal which is finally decoded from the pulse train; a pulse-coded signal will tolerate a much greater level of noise and other ill-treatment than a directly transmitted signal will before the end-product becomes useless.

This method of coding a signal into a series of pulses is known as pulse code modulation (pcm), and the method of generating and decoding these signals is described below.

* 21 Rushes Road, Petersfield, Hants GU32 3BW.

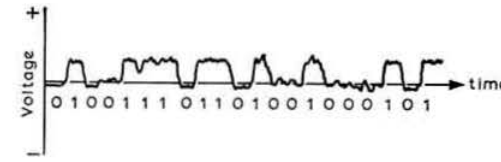

Fig 2. The pulse train of Fig 1 with attenuation, noise and distortion after transmission over an imperfect signal path

Sampling and bandwidth

A signal such as that represented by the waveform of Fig 3(a) may be partly represented by measuring its amplitude at intervals and recording the successive amplitudes. This process is called sampling and if the samples (measurements) are taken at regular intervals the number per second is called the sampling rate. It may be shown mathematically that if the signal contains no components with frequencies greater than half the sampling rate, no information will be lost by representing the signal by a series of samples as described. In other words, the original signal may be perfectly reconstructed from the samples, provided that the sampling rate is greater than twice the highest frequency contained in the signal. Human speech can be understood sufficiently well for communication purposes if all the frequency components above 3·4kHz are omitted, therefore speech may be adequately represented by taking samples at a rate of 8,000/s which is the standard sampling rate for communications purposes.

Quantization and coding

Supposing that the signal to be transmitted is speech and that it is sampled at the points marked "s" in Fig 3(a) separated by 125μs (8,000 samples/s). The sample amplitudes could be transmitted, and hence all the information required to reconstruct an understandable version of the original signal, as a series of brief pulses (Fig 3(b)). The amplitude of each pulse would be proportional to that of the corresponding

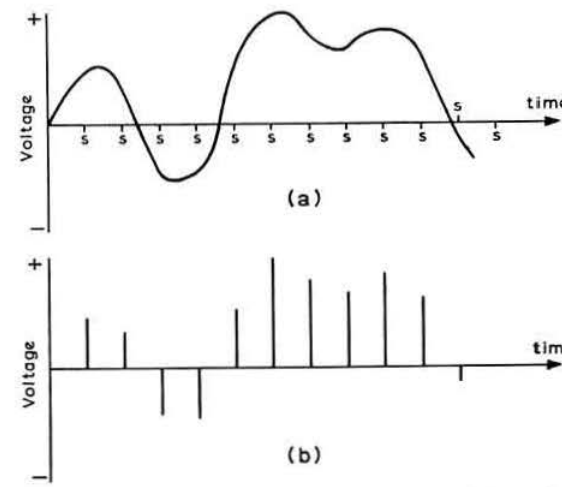

Fig 3. (a) A possible signal waveform, sampling points marked "s". (b) A train of brief, variable-amplitude pulses representing the sample amplitudes in (a)

sample and the original signal could be recovered at the receiving end simply by passing the train of pulses through a low-pass filter. However, this would not be an improvement over direct transmission of the signal waveform; in fact, it would be a step backwards, since the bandwidth required for transmission would be greater. What is needed is a way of representing each sample by a series of pulses of equal amplitude rather than by a single pulse of variable amplitude in order to obtain the advantages of pulse transmission mentioned earlier.

The simplest way to do this is to assign to each possible sample amplitude a binary number and to transmit that number as a series of pulses whenever that particular sample amplitude arises. The pulse train corresponding to the binary number 11010, which corresponds to 26 in decimal numbers, is shown in Fig 4. However, in the system described so far the sample amplitude is a continuous variable, ie it can take an infinite number of possible values within the range of the sample signal amplitude. Clearly an infinite variety of binary numbers is unacceptable, since some of them would then be infinitely long, so the number of values which the sample amplitude may take must be restricted. This process is called quantization and could be carried out in a practical system as follows.

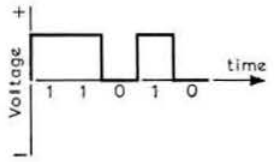

Fig 4. Pulse train corresponding to the binary number 11010

Each sample amplitude is compared with a set of "standard" levels and the nearest of these standard levels is taken as an approximation to the sample amplitude and assumed to be equal to the sample amplitude. After quantization, the sample amplitude must thus be one of a restricted repertoire of amplitudes and the number of possible amplitudes is, in effect, limited. The mechanism of quantization, therefore, introduces small errors into the samples, since each quantized sample is really only an approximation to the true sample, and these distortions manifest themselves as noise in the copy of the original signal reproduced at the receiver. This is known as quantization noise and its amplitude is dependent on the number of quantization levels (the "standard" levels mentioned above) and on their spacings.

The quantization levels in the hypothetical system could be equally spaced (linear quantization) but there are advantages, at least in the transmission of speech or music, to be gained from a non-linear quantization process in which the quantization levels are closer together at low amplitudes than at high ones. This is because in speech small amplitudes are far more common than large ones and, in fact, a logarithmic distribution of quantization levels gives the smallest distortion of small signals and, hence, the best overall noise performance. Clearly the number of quantization levels employed greatly affects the performance, since the more levels there are the better will be the approximations made in the quantization process. In practice, telephone quality speech may be transmitted with 128 ($=2^7$) logarithmically spaced levels (64 positive and 64 negative), and 8,192 ($=2^{13}$) levels are adequate for broadcast quality sound.

After quantization, generation of the binary number corresponding to each sample is needed, and this is done by

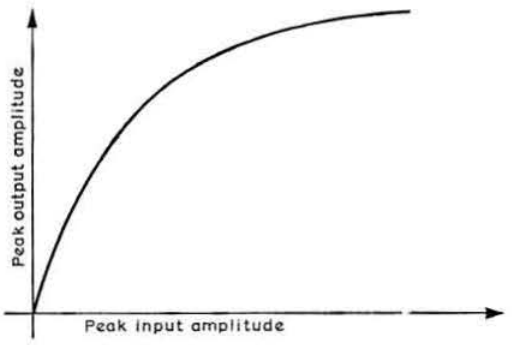

Fig 5. Characteristics of a signal compressor

an analogue digital converter. Successive binary numbers are converted into pulse trains, synchronizing pulses are added and the pulses are transmitted. Two points must be made. First, although the quantization and coding processes have been presented separately above, the two are usually performed by the same circuit, the analogue-digital converter. Second, the analogue-digital converter usually results in linear quantization and so effectively non-linear quantization is achieved by passing the samples through a logarithmic compressor which has the property of amplifying large signals less than small ones, as shown in Fig 5, and therefore of "compressing" the range of sample amplitudes.

A practical system and tdm

That, then, is the outline of a generalized pcm system, as illustrated in Fig 6. Now let us consider a practical system; that employed by the British Post Office for trunk telephone links in some urban areas. The sampling rate is 8kHz and a logarithmic quantization system with 128 levels (corresponding to seven digit binary numbers) is employed, which yields a signal to quantization-noise level of at least 25dB. Each sample is represented by seven binary digits (bits) but an eighth bit is added for the purpose of synchronizing the receiver decoding equipment at the other end of the trunk cable.

One of the advantages of pcm over other systems is that more than one speech signal may be transmitted in coded form over a single pair of wires. This is done by transmitting the pulse train corresponding to each sample in less time than the sampling period and transmitting pulse trains corresponding to samples from other signals during the

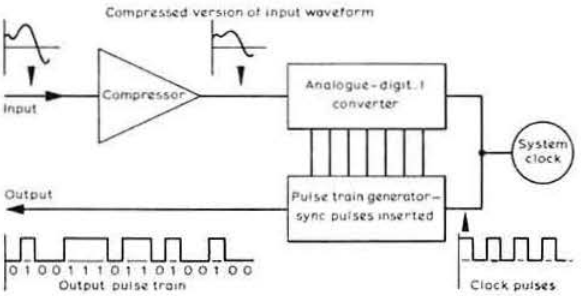
Compressed version of input waveform
Input
Compressor
Analogue-digit..l converter
System clock
Output
Pulse train generator– sync pulses inserted
0 1 0 0 1 1 1 0 1 1 0 1 0 0 1 0 0
Output pulse train
Clock pulses

Fig 6. Block diagram of a general pcm system

spare time thus made available. This is called time division multiplexing (tdm) and is employed in the British Post Office pcm system to transmit 24 speech channels over a single pair of wires.

Returning to actual details of the Post Office system; a group of 24 coded samples (one from each channel) is called a frame and is transmitted in 125μs (one sampling period). Thus each sample must be transmitted in about 5·2μs, this interval is known as the channel time slot. Each sample is transmitted as a sequence of eight pulses (or rather as a sequence of eight intervals during each of which a pulse may or may not be present) including the synchronization digit. Each pulse must therefore be transmitted in 0·65μs or less. In practice, only 0·325μs of this time is allocated to pulse transmission, the remaining half being an unoccupied guard period to allow regenerators and receiving equipment to recover from the last pulse before receiving the next one. Binary zeros are represented by negative voltage levels and binary ones by positive levels, and the pulse train is then passed through a process known as output bipolarization which has the effect of reversing the polarity of every alternate pulse, see Fig 7. Bipolarization removes the dc component from the signal (pulse train) and halves the effective bandwidth needed for transmission.

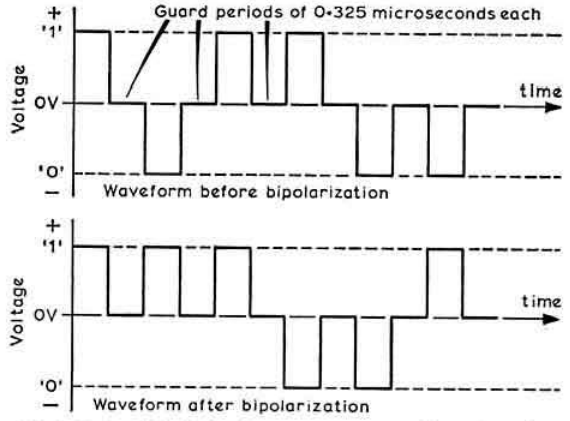

Fig 7. Output bipolarization reverses every alternate pulse

Receiving and decoding

The advantages of pulse-coded transmission from the viewpoint of immunity to random noise have already been outlined. Practical immunity from low-amplitude noise can be obtained at the receiving end simply by passing the received signal through trigger circuits as shown in Fig 8.

In practical systems rather more sophisticated methods are often employed. The actual mechanism of decoding is, at least in principle, fairly simple; the synchronizing information is extracted from the signal and used to "gate" groups of seven bits each into a register consisting of bistable circuits which functions as a digital-analogue converter and recovers the original sample amplitudes one at a time. The samples are then expanded (expansion is the reverse process of compression) to recover their original amplitude distribution and de-multiplexed into their respective channels by what amounts to a fast switching process synchronized to the multiplexing circuits in the coding equipment. In each channel the stream of samples is passed through a low-pass filter which removes any components

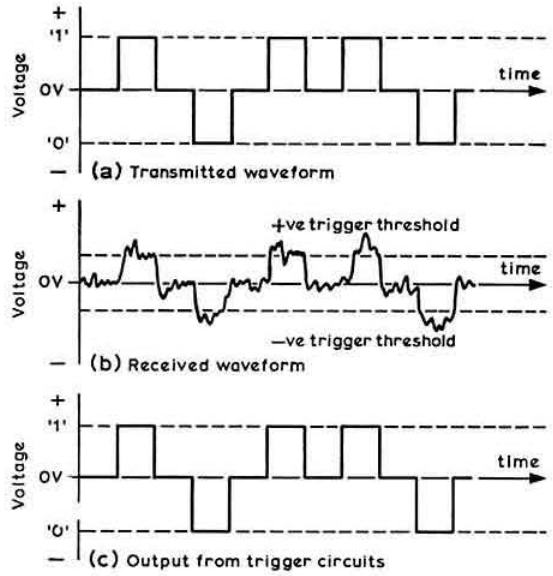

Fig 8. (a) transmitted waveform, (b) received waveform, (c) output from trigger circuits

above half the sampling frequency and the original signal has been recovered with negligible distortion.

Conclusion

Applications of pcm are not restricted to noise-free communications; it can be applied to tape recording to obtain a signal-to-noise ratio far superior to that obtainable with analogue recording methods, and it may be applied to military and other secret communications since it is impossible for anyone to decipher an intercepted pcm signal unless such details as sampling rate, coding system etc are known. Finally, it should be mentioned that pcm is an example of knowledge being ahead of the technology to apply it, for the principles of pcm have been known since it was devised in 1936 by A. H. Reeves, but only recently with the availability of low-cost digital microcircuits has it been a feasible economic proposition.

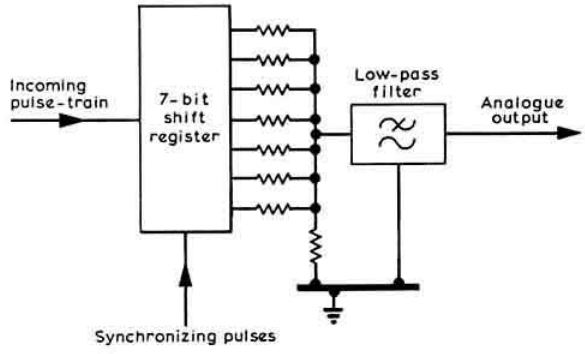

Fig 9. Block diagram of an analogue-to-digital converter. In the matrix of "weighting" resistors, the smallest resistors correspond to the most significant pulses, each pulse being half as significant as the last